Assessing Student Learning in the Exhibit Development Learning Model

Prepared for
the Museum Magnet Elementary School
and the Science Museum of Minnesota

By
Debra Ingram, Ph.D. and Douglas Huffman, Ph.D.
Center for Applied Research and Educational Improvement
College of Education and Human Development
University of Minnesota

August, 1997

Ingram, D. and Huffman, D. (1997). *Assessing Student Learning in the Exhibit Development Learning Model.* Minneapolis, MN: Center for Applied Research and Educational Improvement.

Acknowledgements

Funding for this project was provided by the Institute of Museum Services through its Museum Leadership Initiatives program.

The authors gratefully acknowledge the participation in this project by the following staff at the Museum Magnet Elementary School:

- Kelly Finnerty, Science Museum of Minnesota Program Coordinator.
- Wayne E. Gilman, Teacher.
- Dennis St. Sauver, Principal.
- Karen Thimmesch, Teacher.

Frances P. Lawrenz, Associate Dean of the Graduate School and Professor, Department of Curriculum and Instruction, provided consultation and feedback on the design and implementation of the project.

Jeanette Colby, of the Center for Applied Research and Educational Improvement, assisted in this project by scoring the hands-on performance tasks and preparing the report sections on the testing results.

Table of Contents

List of Figures

EXECUTIVE SUMMARY

This report describes the results of a project to develop and pilot test an assessment system for the exhibit development learning model used at the Museum Magnet Elementary School in St. Paul, Minnesota. The project was funded through the Museum Leadership Initiatives program of the Institute on Museum Services. Staff at the Center for Applied Research and Educational Improvement at the University of Minnesota were hired as consultants by the Science Museum of Minnesota to work on the project with museum educators and school faculty at the Museum Magnet Elementary School.

Project Approach

The project involved three major stages: needs assessment, development of an assessment portfolio, and pilot testing of the portfolio. The specific processes used in each stage are described below.

Needs Assessment. The purpose of this stage was to understand and describe the exhibit development learning model. In order to see first-hand how exhibit-making is incorporated into classroom instruction, the consultants observed classes, the fall grade level exhibits, and planning meetings between teachers and museum staff. The consultants also conducted two focus groups; one group with teachers and a second group with parents of students who had attended the school for several years.

Develop Assessment Set. After developing an understanding of the learning model and generating an initial list of skills students may learn in the model, the consultants worked for several months with the two fourth grade teachers and the program coordinator to develop a process for assessing student learning in exhibit-making. The first step was to finalize the list of skills that fourth grade students learn when they develop exhibits. The next step was to define varying levels of competence for fourth grade students on each skill. Together, the list of skills and the statements describing levels of competence make up what the school calls a rubric.

After the fourth grade rubric was established, the consultants identified possible hands-on performance tasks that could be used to assess these skills. The potential tasks were reviewed by the project team and the final selection was based on those tasks that received the highest total of the individual ratings.

Pilot-Test. The scoring rubric was pilot tested in the spring of 1997 by two fourth grade teachers with the intent of modifying it for school-wide use the following year. The hands-on testing was conducted with students in the two fourth grade classes. The students participated in one of two tests: 1) A series of hands-on stations, or 2) a full investigation.

Rubric for Exhibit Development Process

A copy of the fourth grade rubric developed by the team is shown on pages 7 and 8 of the report. For each of the four stages in the museum process, the rubric defines skills fourth graders are expected to develop and statements that describe three levels of competence for each skill. The statements at the proficient level describe what is expected for students at the fourth grade level. Statements at the excellent level describe a skill beyond what is

expected for students at that grade level. The teachers developed the levels with the intention that only about five percent of the students would be capable of demonstrating a skill at the excellent level.

The rubric is intended to be general enough to apply to all exhibit topics. Although not every skill will be used in every exhibit, it is expected that each skill would be covered in at least one of the exhibits each year. Teachers in other grades may need to modify the list of skills and the statements describing the levels for each skill in order to make the rubric relevant for their students.

Hands-On Testing

To better understand students' exhibit process skills, two different hands-on tests were administered to the students in the two fourth grade classes: 1) A series of hands-on stations, and 2) a full investigation. (See Appendix C for copies of the instruments.) Both tests included items developed by the National Assessment of Educational Progress (NAEP) and were selected to match the type of skills expected of students in the exhibit development learning model.

Overall, students performed relatively well on this set of performance tests. The majority of students were able to successfully use the equipment to solve problems. No differences were found by sex of students, suggesting that the program at the Museum Magnet School is working equally well for both boys and girls. In general, students performed best at making simple observations, conducting simple experiments, and using measuring tools. Students had more difficulty making detailed observations, writing complete explanations, and designing and conducting complex experiments. The results suggest that there are specific exhibit process skill that students still need to develop.

Concluding Remarks

As a result of this project, staff at the Museum Magnet School have established a process for defining and assessing what students learn as they develop exhibits. In the coming year, staff will continue to refine this process as they expand it to all grade levels and use the rubric as a tool for planning instruction and communicating with parents. In addition, the results of the hands-on testing conducted as part of this project provide information for staff to use as they continue to refine the exhibit development learning model. The implications of this project for future use of the exhibit development learning model are described below.

<u>Implications for Strengthening the Exhibit Development Learning Model</u>

The process of developing an assessment system for the exhibit development model has illuminated aspects of the model that need clarification. Future use of the model could be strengthened by clarifying issues such as the following:

- How do staff distinquish what gets taught and assessed in the exhibit development process from what gets taught in other areas of the curriculum? A strength of the model is that it provides a framework for interdisciplinary instruction, but this also makes it difficult to differentiate between what students learn by developing exhibits and what they learn through other classroom activities. There may not be one right answer to this question, but in order for the model to be used consistently

across classrooms, the staff need a shared understanding of how the model will be used at this school.

- What are the distinguishing characteristics of the model, the practices that must occur in order for the model to be implemented effectively? The project results indicate that classroom teachers vary in how they integrate the exhibit process with the academic content they are required to cover. Some variation is necessary given the different needs of students and the different academic requirements at each grade level. In order for the model to have substance, however, there must be a core of activities and skills that distinguish the model from other instructional practices. Some possible distinctions may be that exhibit development always involves the use of real objects and always requires students to explain what they have done and learned to other people.

The hands-on tests administered in the project identified several specific exhibit process skills that students still need to develop. To strengthen their use of the model, staff may want to adjust their instruction to address these issues:

- Observe. According to the exhibit development rubric, proficiency is defined as making accurate, orderly observations. The results of this performance test, however, indicate that many students need development on this skill. Students were able to make simple straightforward observations, but rich, detailed descriptions of observations were rare. For example, on the popcorn station students watched a phenomenon with many different events to observe. When asked to describe what they observed, the majority of students only wrote down one or two observations with very minimal descriptions despite the fact there were at least six different events to observe. These results imply that students should be encouraged to make more careful, detailed observations.

- Design a Test / Conduct a Test. Designing and conducting a test is a complex set of skills and the results of the "Sugar Test" indicate that students need work in this area. One specific skill that needs work is learning how to conduct a "fair" test. On the sugar test, a fair test would involve using equal amounts of water to determine if loose sugar or cube sugar dissolves faster. One should also use equal amounts of sugar and put both types of sugar in the water for equal amount of time. None of the students who took the test explicitly controlled all of these variables. Students should be taught to carefully control variables to insure all tests they conduct are fair.

  Not only is it important to use equal amounts of water, sugar and time, but it is also important to systematically measure each of these variables. Students were supplied with a stopwatch to measure the time and a measuring cup to measure the amount of water, but students very rarely used these instruments. They should be encouraged to use more precise measurement techniques.

  Another important aspect of conducting a test is repeating the test more than once to make sure the results are consistent. None of students conducted the sugar test more than once.

- Explain. Students were also able to provide simple, straightforward explanations, but detailed explanations were lacking in students responses. Students were most likely to merely provide one explanation, rather than elaborate on their explanations. Students should be encouraged to write more detailed, thorough explanations.

Implications for Future Evaluation of the Model

Next steps for evaluating the exhibit development learning model are as follows:

- Develop rubrics for other grade levels. This task was already started at the summer workshops in August.

- Decide how often the hands-on testing should be used and which grade levels should be involved. The relationship between these tests and student ratings on the rubric also needs to be determined. For example, will results of the hands-on test influence teacher ratings of certain skills on the rubric or will the hands-on test results be reported separately?

- To determine the effectiveness of the exhibit development learning model in comparison to traditional instructional practices, administer the hands-on tests to a comparable group of students at a school that uses traditional instructional practices. In order for the comparison to be meaningful, the comparison school must have similar instructional goals.

- This project focused on the exhibit development learning model but there are other characteristics of the Museum Magnet School that could be the focus of future evaluation efforts. For example, future studies could assess the effectiveness of other Museum Magnet School components such as the emphasis on parent involvement, students working with adult mentors (either museum staff, community members, or parents), and students working cooperatively.

This report describes the results of a project to develop and pilot test an assessment system for the exhibit development learning model used at the Museum Magnet Elementary School in St. Paul, Minnesota. The project was funded through the Museum Leadership Initiatives program of the Institute on Museum Services. Staff at the Center for Applied Research and Educational Improvement at the University of Minnesota were hired as consultants by the Science Museum of Minnesota to work on the project with museum educators and school faculty at the Museum Magnet Elementary School.

I. Purpose of Project

Despite a general impression among staff that the exhibit development process is valuable, the school has had difficulty identifying and assessing what students learn during the process. Staff were concerned that the district mandated tests did not provide an accurate reflection of what students were learning at the school because the tests were not designed to measure skills such as critical thinking or creative problem solving. In addition, the school needed to address concerns raised by some parents who thought the exhibit process was a distraction from what their children needed to learn in school. This project is an attempt to address these issues by developing a method for assessing the contributions of the exhibit development process to student learning.

II. Project Approach

The project involved three major stages: needs assessment, development of an assessment portfolio, and pilot testing of the portfolio. The specific processes used in each stage are described below.

Needs Assessment. The purpose of this stage was to understand and describe the exhibit development learning model. In order to see first-hand how exhibit-making is incorporated into classroom instruction, the consultants observed classes, the fall grade level exhibits, and planning meetings between teachers and museum staff. The consultants also conducted two focus groups.

The first focus group was with teachers representing a range of grade levels and specialty areas at the school. Teachers were asked to describe how they use the exhibit development process in their instruction and what skills they think students learn through exhibit-making. The second focus group was with parents of students who had been attending the school for several years. The parents were asked to describe what they thought their children learn through the exhibits.

Develop Assessment Set. After developing an understanding of the learning model and generating an initial list of skills students may learn in the model, the consultants worked for several months with the two fourth grade teachers and the program coordinator to develop a process for assessing student learning in exhibit-making. The first step was to finalize the list of skills that students learn when they develop exhibits. The next step was to define varying levels of competence for each skill. Together, the list of skills and the statements describing levels of competence make up what the school calls a rubric.

After the rubric was established, the consultants identified possible hands-on performance tasks that could be used to assess these skills. The potential tasks were reviewed by the project team and the final selection was based on those tasks that received the highest total of the individual ratings.

Pilot-Test. The scoring rubric was pilot tested in the spring of 1997 by two fourth grade teachers with the intent of modifying it for school-wide use the following year. The hands-on testing was conducted with students in the two fourth grade classes. The students participated in one of two tests: 1) A series of hands-on stations, or 2) a full investigation. For the former, each student rotated between four different stations that assessed students' ability to solve problems, make observations and explain the results of their experimenting. A small sample of students participated in the full investigation test where they were asked to design and conduct a complete experiment. Each of the above assessment activities and the corresponding results are described in more detail later in the report.

III. Description of the Museum Magnet Elementary School

The Museum Magnet Elementary School (MMS), a collaboration of the Science Museum of Minnesota (SMM) and the Saint Paul School District, opened in September, 1991. The school is one of only five in the nation that have a formal partnership with a museum, and this partnership is the longest running.

The school is located approximately one mile from the SMM. The layout includes twelve classrooms clustered around two large exhibit halls and a larger amount of storage space, needed to store exhibit materials and collections, than in a typical school.

The mission of the MMS is to nuture creative thinkers who can work cooperatively to solve problems. The core curriculum is based on the premise that the exhibit development process used by museums can stimulate students' creativity and critical thinking and provide an interdisciplinary framework for instruction.

The partnership with the SMM provides students with an enriched educational environment, through access to both its collections of objects and the expertise of its staff. Students have access to a diversity of objects through several means. First, because the museum donates objects it can no longer use to the MMS, the physical environment of the school contains a greater array of objects than most other schools. Second, the museum also lends objects to the school for special projects, or makes them available to students during field trips to the museum. Finally, access to the expertise of museum staff also contributes to the enriched environment of the school. In addition to the museum staff that are located on-site (described below), other museum staff contribute their expertise to the exhibit development process by working with the students at the museum or by serving as guest instructors at the school.

The museum has three staff people located on-site: a full-time program coordinator, a part-time science instructor, and a part-time assistant coordinator. The museum staff works closely with the school's teaching and administrative staff. Program decisions are made jointly by an administrative team of school and museum personnel with input from parents and students.

During the 1996-1997 school year, 372 students were enrolled at the school in grades K-6. As a magnet school in a district with school choice, the school draws students from throughout the city. The student body is economically and culturally diverse with high representation of groups identified as under-served. Students of color, primarily African-American with a smaller group of Hmong, comprise 57% of the student body. Sixty percent of the students are low income, receiving subsidized school lunches.

In addition to the unique focus on exhibit-making, the school also emphasizes parent involvement, the use of adult mentors (either museum staff, community experts, or parents

with knowledge in a particular area), and students working in cooperative groups. The focus of this project, however, was the exhibit development learning model and how to assess what students learn during the process.

IV. Description of the Exhibit Development Learning Model

Over the past six years, museum staff and staff at the Museum Magnet have worked collaboratively to integrate the process museums use to develop exhibits with the district's learner outcomes in academic content areas. The result of this collaboration is the exhibit development learning model. The purpose of this section is to provide first a general description of this model, and then illustrations of how the model was used during the 1996-1997 school year. This description is based on information gathered from the following sources: a review of school documents, including a paper published by the museum program coordinator; focus groups; and observations of meetings, classes, field trips and exhibits.

Conceptual Model

In their early work to develop the exhibit development learning model, museum and school staff established broad goals for the program, and learner outcomes at each grade level. The five broad goals for the program are as follows:

1. Learners will use museum artifacts, specimens, exhibits, educational programming and staff to support inquiry learning.

2. Through the museum process learners will use objects and collections as tools for research.

3. Learners will work cooperatively to develop exhibit projects that demonstrate their critical thinking and creative problem solving skills applied to a theme that integrates academic content in the major subject areas.

4. Learners will interpret their exhibit projects for an audience through a variety of written and verbal modes including writing labels and scripts, performing demonstrations and dramatizations and leading tours and visitor activities.

5. Through exposure to many museums and in-depth knowledge of SMM in particular, learners will be involved in the continual creation of a museum by kids for kids with the school environment.

Staff also developed the Museum Process Wheel to structure the exhibit development process for students. The model is still being refined as the school gains experience with its implementation, but in general, the model includes four basic steps:

Explore - students investigate a topic by making observations, researching, and demonstrating a willingness to see things in new ways.

Experiment - students pose questions of "why" and "what if" and conduct tests to answer questions.

Explain - students describe the results of their explorations and experimenting by working collaboratively to identify patterns and make new connections.

Exhibit - new understandings are presented to the school community in a variety of formats, using multiple communication media and tools.

The exhibit development learning model is an attempt to blend the informal learning environment of a museum with the formal learning environment of an elementary school. Through the museum process, students gain experience with a wide range of actual objects and learn how to explain to other people what they have done and what they have learned. School staff describe the difference between student exhibits and typical student projects as follows:

> An exhibit may include a project (something that a student can take home), but it is more than that. An exhibit requires students to interpret their project for other people. They have to be able to tell you what they did, and how they did it. During the exhibit they have to explain the objects or content of the exhibit, and show the process of developing the exhibit. An exhibit is less structured than a typical science fair project where there is always a hypothesis, an experiment, and a conclusion.

School staff believe that students are more likely to remember what they have learned if they have to explain it to someone else and answer someone's questions about it.

The use of objects, rather than pictures or descriptions of objects, is also a key part of the learning model. Staff believe that students get more exited about learning when they use real objects, and that some students may learn better by being able to interact with artifacts.

Parents involved in focus groups during the needs assessment phase of this project generally confirmed this belief. They were enthusiastic about the model and believe that it is dramatically helping their students become more observant and inquisitive. One parent said that his children were more likely to ask questions about things they encountered in their environment; another said her children were more interested in visiting the library to get information about questions they had raised.

The focus group of teachers, also conducted during the needs assessment phase, generated an extensive list of skills that students learn during exhibit making. The skills range from traditional education topics such as reading and writing, to skills needed for working in groups, such as listening, negotiating, and cooperating. The skills identified by both the parent group and the teacher group were used as a starting point for the team to develop the rubric. The complete list of skills described by each group is shown in the Appendix.

Using the Model

Each year, the first step in the exhibit development process is the selection of exhibit topics. For the fall exhibits, each grade level selects its own topic. For the spring exhibits, one theme is selected for the entire school. During the 1996-1997 school year, the theme for the spring exhibit was inventions. Topics for grade level exhibits in the fall included Minnesota history in the fourth grade and worms in the first grade.

Once an exhibit topic is selected, the museum's on-site staff meet with teachers at each grade level for a half-day to determine how museum resources can be used in the classroom as students develop exhibits. Potential resources include the following: classroom instruction by on-site museum staff; material resources owned by the museum; field trips to the SMM or other local museums; and pre-visit class sessions conducted by on-site

museum staff to prepare students for field trips. The types of resources used may vary from grade level to grade level and from topic to topic.

Teachers at the MMS integrate the museum process and the academic content they are required to cover in a variety of ways. The fourth grade teachers try to include as much academic content as possible in the exhibit development process in order to cover both the exhibit process and the curriculum content required by the district or the school. For example, the fall exhibit on Minnesota enabled the teachers to incorporate the district requirement that students learn Minnesota history. In the primary grades, where the focus is more on skill development than on covering content, the majority of instruction may be centered around the exhibit topic. For example, for the fall exhibit on worms, students could learn reading by reading about worms.

The fourth-grade exhibit was titled, "Our Minnesota." On-site staff provided activities to prepare students for field trips to the history museum and the SMM, and taught a class session on how to consider the audience for an exhibit when designing the exhibit. The exhibit included information on history, geography, people, industry, and natural resources. Students investigated various Minnesota related topics such as logging, lumberjacks, rocks & minerals, iron ore mining, and artifacts from native Minnesotans. The students worked in small groups to prepare exhibits on their topics. In the first step, students explored their topic by gathering information from books and other media sources. This topic did not lend itself to much experimenting, so students focused most of their effort on explaining and exhibiting the information they had gathered about the state of Minnesota. Students explained what they had learned through written words, pictures, maps, and dramatic plays of the information they had gathered. On exhibit day, small groups of four to five students presented their work to a large group of parents, students, and teachers. One group used a talk show format to present the information they had learned about lumberjacks. After the formal presentations, the audience wandered around the various physical exhibits the students had created. Displays included artifacts, interactive games, and poster boards with maps and drawings. The students even published a newspaper on Minnesota that included poetry, songs and essays.

In preparation for the spring all-school exhibit on inventions, on-site museum staff taught six classes at each grade level. These classes focused on getting students to recognize problems and then to invent a solution for the problem. The first activity in the series had students improve the design of pencils; the third activity had them solve a problem in their own life and develop a model for a solution. In one of the classes that was observed students brainstormed activities they enjoy such as skiing, biking, and reading a book. Students were then asked to brainstorm problems they have experienced with the activity, followed by possible solutions to the problem. Through this process students were able to come up with several possible inventions. A similar process was used to explore possible inventions related to things students dislike doing. Class time was also used to have students begin building their inventions with everyday materials such as egg carton, cardboard boxes and plastic bottles.

The spring exhibit included exhibits in individual classrooms, a play by the sixth grade, and a song by the first grade. Each of these activities was centered on the theme of inventions as problem solving. For example, the play demonstrated that inventors see problems as opportunities. In a second grade classroom, an exhibit showed the different uses for peanuts that students had invented. The fourth grade invented "mini-mini" golf courses and the third grade invented cookies. In addition, a design team of 20 students drawn from grades two through six worked with museum staff to develop the "Fabulous Rube Goldberg Marble Machine." The group built a complex device to carry a marble from

one side of the exhibit hall to another. In order for the marble to make it to its destination, the students needed to work together, each carrying out their task at the right time.

V. Rubric for Exhibit Development Process

On the current report card used by the MMS there are several sections for recording grades; one of these sections, labeled "museum process," is reserved for reporting on student learning related to exhibit development. Teachers vary in how they decide what grade to give a student on this portion of the report. The first grade bases the grade on a portfolio of student work; one of the fourth grade teachers uses a checklist of skills the students should be able to demonstrate. Other teachers acknowledge that their decision of what grade to place in this report card section is totally subjective and is based more on effort than achievement.

The original intention for this project was to develop a set of authentic assessments, a portfolio, that would measure skills students learn during exhibit-making. The project focused on developing assessments for the fourth grade, with the assumption that the resulting product could be modified for use by other grades in the following year. A team, comprised of two fourth grade teachers, the museum program coordinator, and the two consultants from the University of Minnesota worked to first identify skills that fourth graders learn during the exhibit process and then to design a method for assessing these skills. A first grade teacher also participated in early meetings of the project to ensure that what was being developed was consistent with how the exhibit process is used in the primary grades. During the process of development the concept of a portfolio was dropped to make the assessment more applicable to a range of classrooms. The end result is a rubric, made up of a list of skills and defined levels of competence for these skills, and a set of hands-on performance tasks related to these skills.

A copy of the fourth grade rubric developed by the team is shown on the next page. For each of the four stages in the museum process, the rubric defines skills fourth graders are expected to develop and statements that describe three levels of competence for each skill. The statements at the proficient level describe what is expected for students at the fourth grade level. Statements at the excellent level describe a skill beyond what is expected for students at that grade level. The teachers developed the levels with the intention that only about five percent of the students would be capable of demonstrating a skill at the excellent level.

The rubric is intended to be general enough to apply to all exhibit topics. Although not every skill will be used in every exhibit, it is expected that each skill would be covered in at least one of the exhibits each year. Teachers in other grades may need to modify the list of skills and the statements describing the levels for each skill in order to make the rubric relevant for their students.

Figure 1: Rubric for Exhibit Development Process

SKILL	LEVEL	
EXPLORE		
Brainstorm •Make a web or list of ideas	ND P E	Student can come up with fewer than 4 ideas on this topic on their own. Student can come up with 4-6 ideas related to this topic on their own. Student can come up with more than 6 ideas related to this topic on their own.
Form questions •Write or ask a researchable question	ND P E	Question could be answered in one sentence. Requires basic information. Question could be answered in a paragraph; requires explaining and comparing. Question is interesting and engaging and may have conflicting answers. Requires a 3 paragraph essay or more to answer.
Observe real objects or demonstrations •Draw •Tell •Write	ND P E	Inaccurate, unorganized, minimal description. Accurate documentation of observation; organizes and records information in an orderly manner. Presents a rich and detailed description of the observation through two of the three options (draw, write, tell).
Research (gather information) •Use resources to answer researchable questions	ND P E	Uses less than 3 different sources. Records disorganized and information is unrelated to question. Uses at least 3 different sources (people, technology, books, found objects, field trips, and artifacts) to appropriately search for information. Gathers relevant information. Records information in an orderly manner. Uses more than 3 sources. Records related information using an outline, list, or web.
EXPERIMENT		
Use measurement tools •Select and use measurement tools accurately and appropriately	ND P E	After lessons in use, student continues to need help in using tool appropriately or accurately. After lessons in use, student can select tool and use it accurately. After lessons, student can select and use tool accurately. Student can also demonstrate and help others in the use of the tool.
Design a test-conduct a test •Try it out	ND P E	Student is unable to follow given directions to conduct a test or experiment without adult assistance. Student is able to follow given directions and complete a test with little adult assistance. Student is able to do test on own, find own sources, try experiments at home, and/or request materials to conduct tests and experiments not required.
•Develop a model	ND P E	After demonstration and examples are shown, student is not able to develop own model without adult direction. After demonstration and examples are shown, student is able to develop a model using own data. Student is able to produce own model and is able to choose and use a different model than the one demonstrated if it is more appropriate.
•Improve, try it out again (models, tests, experiments)	ND P E	Student does not improve or try again. Student makes improvements and adjustments as work is done to complete the task. Student makes improvements and adjustments; then tries a new way and compares the new results to see which was better.
•Record results	ND P E	Does not record findings or record is unuseable. Makes short notes with enough information to know what the notes refer to when read later. Makes identifiable notes, organizes them into a table, chart, graph, or outline.

Rubric for Exhibit Development Process

SKILL	LEVEL	
EXPLAIN		
Tell and teach	ND	Disorganized or inaccurate, copies words.
	P	Uses own words, exhibit content is accurate and organized.
	E	Is proficient and engaging and makes connections to other areas. Can teach in multiple ways.
Use information	E	Uses at least 3 sources of evidence such as: written sources of knowledge, laboratory experiences, observations, and interviews.
	P	Uses 1-2 sources of evidence to support telling.
	ND	Uses no sources to support telling.
EXHIBIT		
Speak	ND	Distracting gestures such as: fidgeting, back to audience, talking to screen and unclear speaking.
	P	Must be heard and enunciated clearly, looks at the audience. Uses audio-visual equipment and microphone properly.
	E	Also engages the audience and uses expression.
Write	ND	Incomplete sentences, unorganized, inappropriate use of grammar.
	P	Complete sentences, appropriate use of grammar, organizes in paragraphs.
	E	Uses adjectives and adverbs; has beginning, middle and end.
Visual presentation •Image selection	ND	Image has minimal relationship to subject matter. Disorganized presentation. Little use of art elements.
	P	Balanced composition. Art elements (line, shape, color, form, texture) convey message. Image relates to subject matter. Text is included in image design.
	E	Art elements convey message in an attractive manner. Original, includes own ideas. Image enhances subject matter.
Productive group interaction	ND	Doesn't compromise or listen, unfocused, non-participation.
	P	Accepts other's suggestions, compromises, willing to share ideas and work.
	E	Also facilitates, encourages others, contributes positively to task completion.
Demonstrate or perform	ND	Ineffective choice or use of props. Unaware of audience. Unclear. Performs with limited expression.
	P	Selects and uses props appropriate for demonstration. Involves audience. Demonstration shows how something is done. Performance includes dramatic interpretation.
	E	Props and activities are interesting. Suggest their own demonstration that relates to the exhibit. Demonstration appears well rehearsed. Performance includes original interpretation.

In addition to using the rubric to assess student performance, it is also intended to be used as a planning guide for teachers. Instead of choosing an exhibit topic and instructional activities based on what looks appealing, teachers can start with the skills listed on the rubric and select an exhibit topic and activities to correspond with these skills. The rubric can also be used to inform students and parents about what skills students will be expected to learn during the process.

During the development of the spring exhibit, the fourth grade teachers pilot-tested the rubric. They also used it at parent conferences in the spring to discuss student performance. Teachers found that the rubric provided parents with a detailed view of their child's performance and identified specific areas in which parents could help their child. The pilot-test also identified several areas where the distinctions between levels of competence needed to be clarified. The teachers will make these adjustments before they use the rubric again in the fall of 1997.

At the MM's summer training in late August, teachers in the other grade levels began to modify the rubric for use in their own classrooms. Because teachers vary in how they integrate the exhibit development process with the academic content they are required to teach, this process provided several challenges.

> Teachers in specialty areas such as music or physical education needed to use a different approach in developing a rubric than classroom teachers who work with one grade level of students all year. For example, the physical education teacher decided to focus on only a few skill areas that were relevant to her work with students, but she developed statements that could be used to define competence in these skills across all grade levels.
>
> Staff questioned whether it is reasonable to expect that students will increase their competence in each skill at every grade level. For example, is the change in writing or speaking skills between fourth and fifth grade significant enough to warrant different descriptions of competency? One option staff may consider is having a rubric for the primary grades and a rubric for the intermediate grades, rather than a different rubric for each grade level.
>
> How do staff distinguish between what gets taught and assessed in the exhibit development process from what gets taught and assessed in other areas such as reading or mathematics? If a student creates a poster, or does a presentation on something that is not related to the exhibit topic, is this considered part of the exhibit process because these skills are listed on the rubric? These questions illustrate the need for the school to further define the model and how it is used in the classroom.

The staff will attempt to resolve these issues in order to have a version of the rubric ready to pilot test at each grade level during the fall exhibit.

In addition to developing the rubric for teacher ratings of student skills, two hands-on assessment activities were selected to correspond with skills listed on the rubric. The next section describes the assessments and the results of this testing.

VI. Hands-On Testing

To better understand students' exhibit process skills, two different hands-on tests were administered to the students in the two fourth grade classes: 1) A series of hands-on stations, and 2) a full investigation. (See Appendix C for copies of the instruments.) Both tests included items developed by the National Assessment of Educational Progress (NAEP) and were selected to match the type of skills expected of students in the exhibit development learning model.

The results of each test are described below and further detail is shown in tables located in the Appendix. Overall, students performed relatively well on this set of performance tests. The majority of students were able to successfully use the equipment to solve problems. No differences were found by sex of students, suggesting that the program at the Museum Magnet School is working equally well for both boys and girls. In general, students performed best at making simple observations, conducting simple experiments, and using measuring tools. Students had more difficulty making detailed observations, writing complete explanations, and designing and conducting complex experiments. The results suggest that there are specific exhibit process skill that students still need to develop.

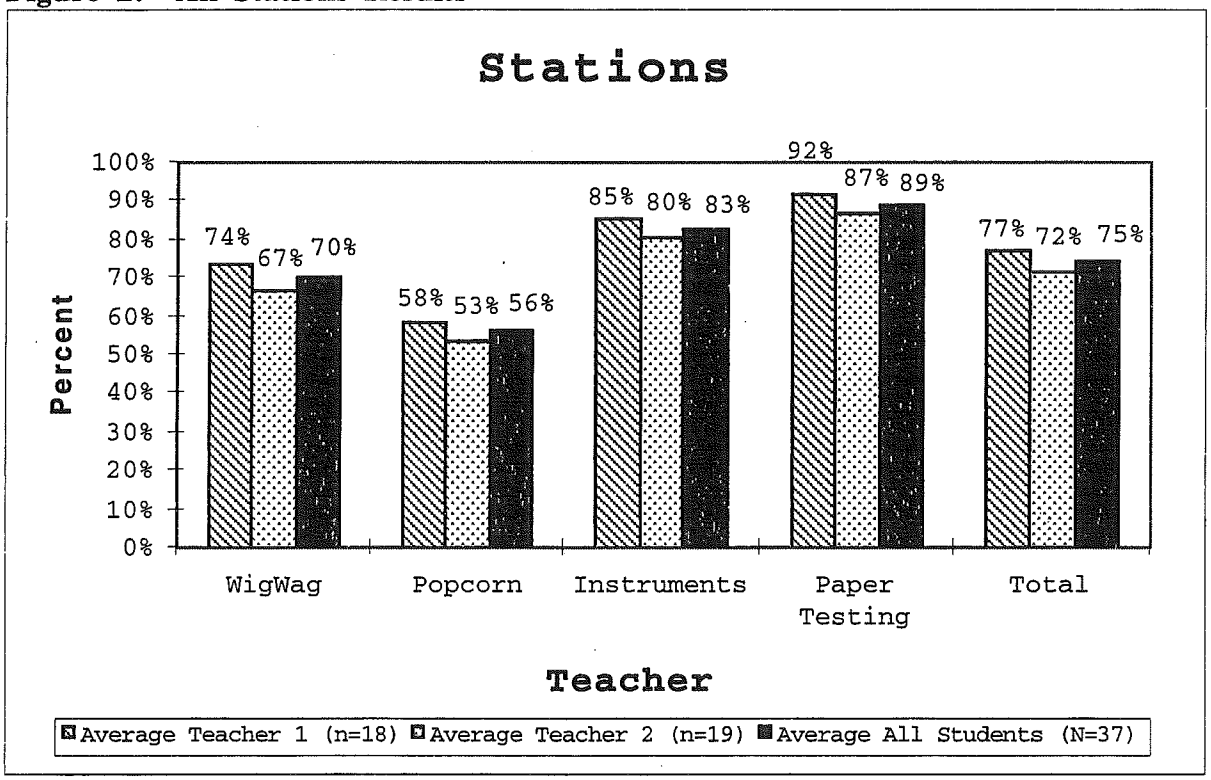
Hands-on Stations

The hands-on stations included four short stations where students were required to use a variety of equipment to answer questions. Each station was selected to focus specifically on exhibit process skills. At the first station, students observed containers of differing weights, placed them on a balancing scale (called a "wigwag"), and determined the relationship between the containers' weights and the movement of the wigwag. This station focused on the skills of observing, making inferences and explaining results. The second station (Popcorn) invited students to drop popcorn into a cup containing water and Alka-Seltzer, observe the result, make inferences, and offer explanations for what they saw. At the third station (Instruments), students were assessed on their ability to use various measuring tools. At station four (Paper Testing), students were asked to drop water onto three types of paper and observe the result, then to predict and explain the result of the same test performed on a fourth kind of paper. This station focused on the skills of experimenting, observing, making inferences, and explaining.

Three different scores were calculated: 1) A total score for all four stations, 2) a score for each individual station, and 3) sub-scores for each of the questions that comprised each station. A total of thirty-seven students participated in this assessment. Statistical comparisons were made between the two classes and between male and female students.

All Stations:
Figure 1 shows scores for all the stations and the total scores for the test as a whole. Overall, students earned 75% of the points possible on this test. The highest scores were on the paper testing station, followed closely by the measurement station. The lowest scores were on the popcorn and wigwag stations. At the paper testing station, students were asked to make one observation about a relatively uncomplicated situation and to give a fairly straightforward explanation and prediction - on these tasks students performed very well. The popcorn and wigwag stations, however, asked students to make multiple observations about complex situations and provide detailed explanations for the phenomena they observed. Lower scores at these stations resulted.

Figure 2: All Stations Results

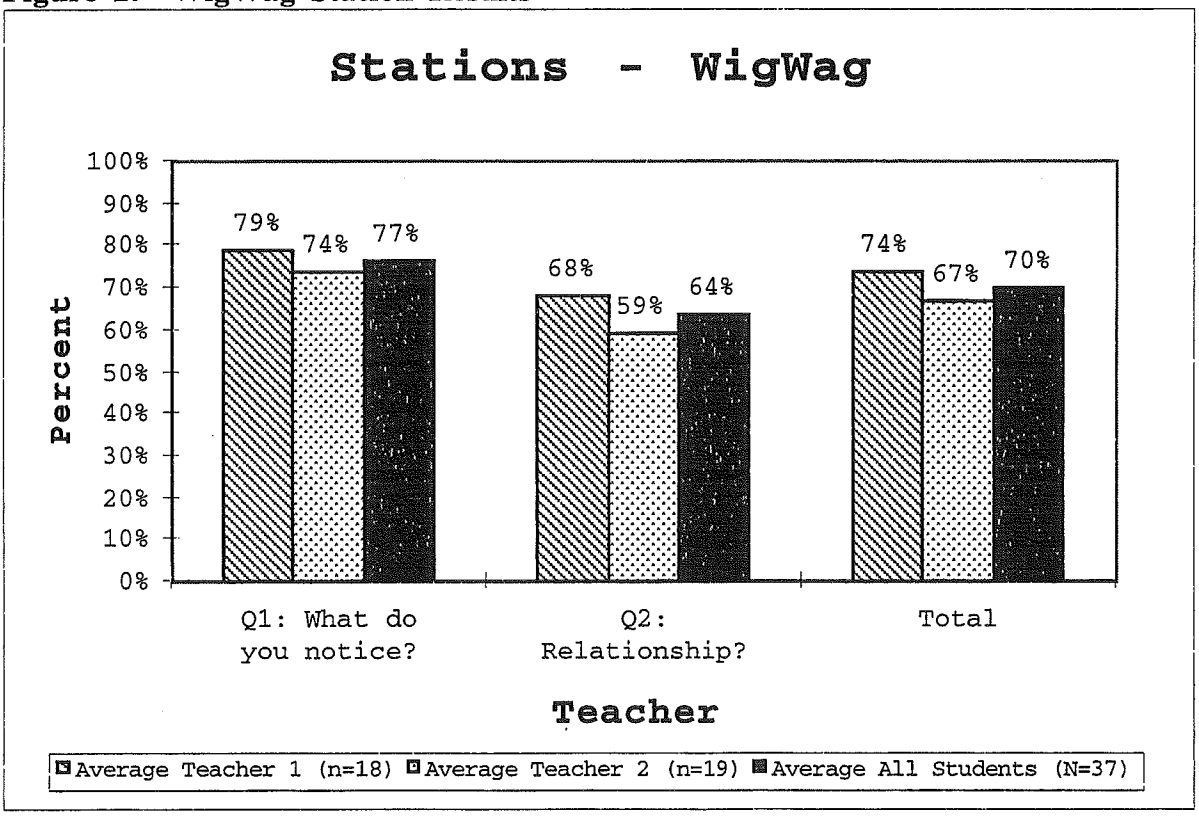


Stations #1: Wigwag

Figure 2 shows the results of the wigwag station. Overall, students earned 70% of the points possible at this station. Question one asked students: " Lift one container at a time. What do you notice?" In order to score the maximum number of points (three points), students were expected to go beyond pointing out that the containers have different weights; they needed to specify how the weighted containers related to each other (container A is the lightest, B is in the middle and C is the heaviest). Sixteen students (43%) received the maximum number of points and another 16 said only that the container had different weights (scoring two of three points possible). Only five students (14%) offered incorrect observations. The average score for the group was 77%.

Question two asked "What is the relationship between the weight of the containers and how the wigwag moves?" The student could earn the maximum number of points (three points) by specifying that the heavier the container, the slower (or lower) the wigwag moves. Two fewer students (14, or 38%) obtained the maximum. (Eight students got the maximum number of points on both questions.) More students (11, or 30%) gave incorrect responses to question two, scoring one point on a three-point scale. Common problems were a lack of focus on the issue of relationship between the containers and the wigwag, repetition of the answer to question one, failure to differentiate among the containers, and incorrect procedure. The average score on question 2 was 64%.

Figure 2: WigWag Station Results



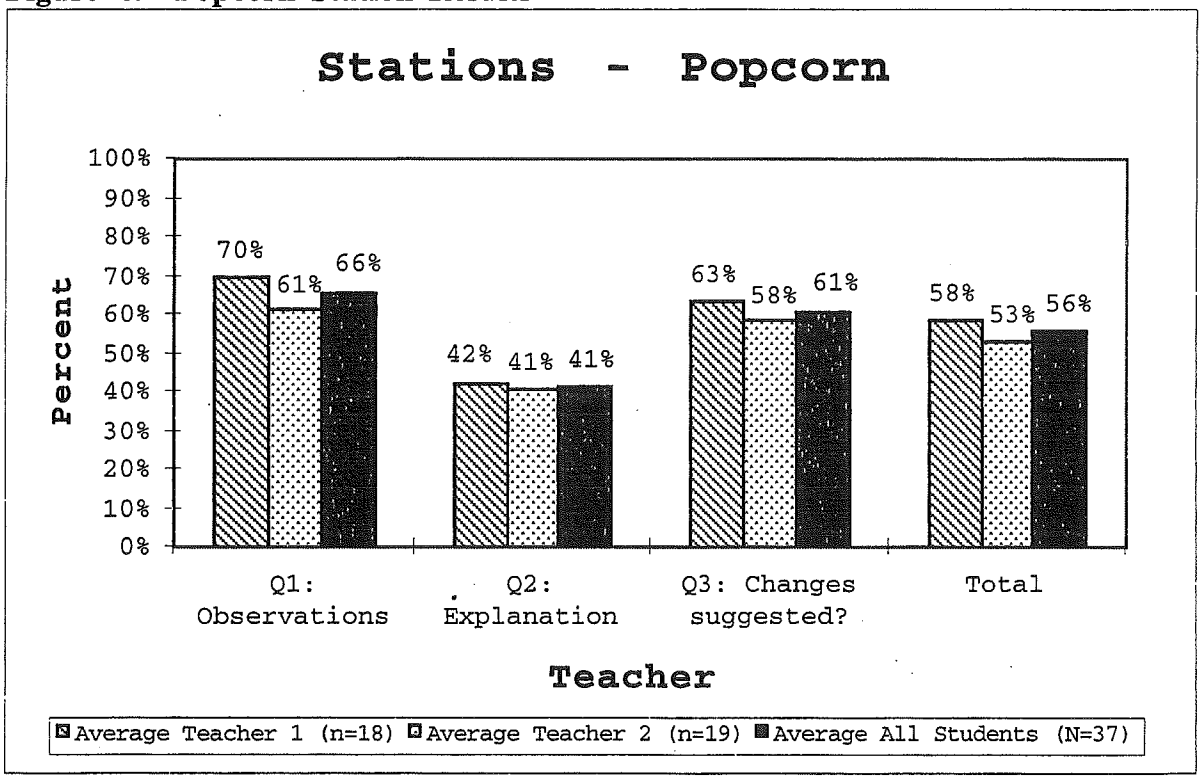
Station #2: Popcorn

Figure 3 illustrates the results of the popcorn station. Overall, students earned 56% of the points possible at this station. Question one asked students, "Describe what you observe" when they dropped an Alka-Seltzer tablet into a cup containing water and popcorn kernels. Students earned the maximum number of points (four) by giving three or more relevant observations. Most students offered only one or two relevant observations (14 and 17 students, respectively) scoring either two or three points. On average, the group earned 66% of the possible points.

In question two, students were asked to give an explanation for the movement of the popcorn in the cup. In order to score the maximum number of points (three), an explicit, complete explanation of the phenomenon was required; none of the students did so. Eleven students gave reasonable explanations (scoring two points on a three-point scale, 67%). Twenty-four gave incorrect responses or offered an observation that would have been appropriate for question one. Two did not respond. The average score was 41%.

Question three asked students to "Name one thing you would change to learn more about what is happening to the popcorn." If students suggested a change or modification of the original conditions they received two points on a two-point scale. Eleven students (30%) did so. Twenty-three (62%) gave incorrect or irrelevant responses (scoring 1 point). Three did not respond. The group average was 61%.

Figure 4: Popcorn Station Results



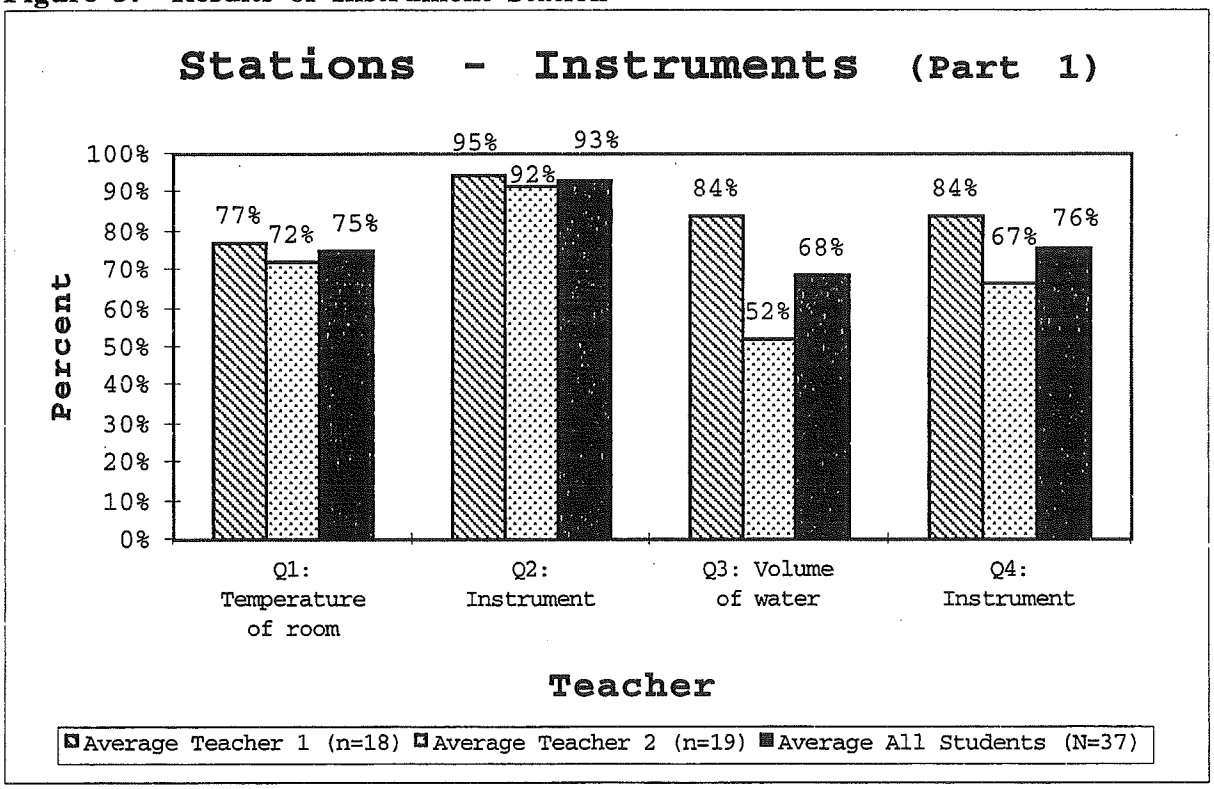
Station #3: Instruments

Figures 4 and 5 show the results of the measuring instruments station. Students were asked to measure something and write down the result, then indicate what instrument they used to conduct the measurement. Overall, students earned 83% of the points possible at this station.

Questions one and two concerned the temperature of the room. Generally, students either responded correctly within plus or minus two degrees Fahrenheit (scoring three on a three-point scale, or 100%), or they responded completely incorrectly (scoring one on a three-point scale, or 33%). Twenty-two students (59%) responded correctly and 13 (35%) responded incorrectly. Two students gave responses within a range of plus or minus four degrees Fahrenheit. The average for the group was 75%. Nearly everyone (93%) used the correct instrument, the thermometer, although few knew how to spell the word and some drew a picture which indicates that they either did not know the name or did not want to attempt the spelling. Full credit was given for these responses.

Questions three and four relate to the measurement of the amount of water in a bottle. In general, students found this measurement the most challenging of the four. Twenty students (54%) accurately measured the volume of water. Of the 16 (43%) that did not, most used the wrong instrument, a ruler rather than a measuring cup. Overall the students in Teacher 1's class scored higher than the students in Teacher 2's class on the two parts of this question, yielding a statistically significant difference a the .05 level.

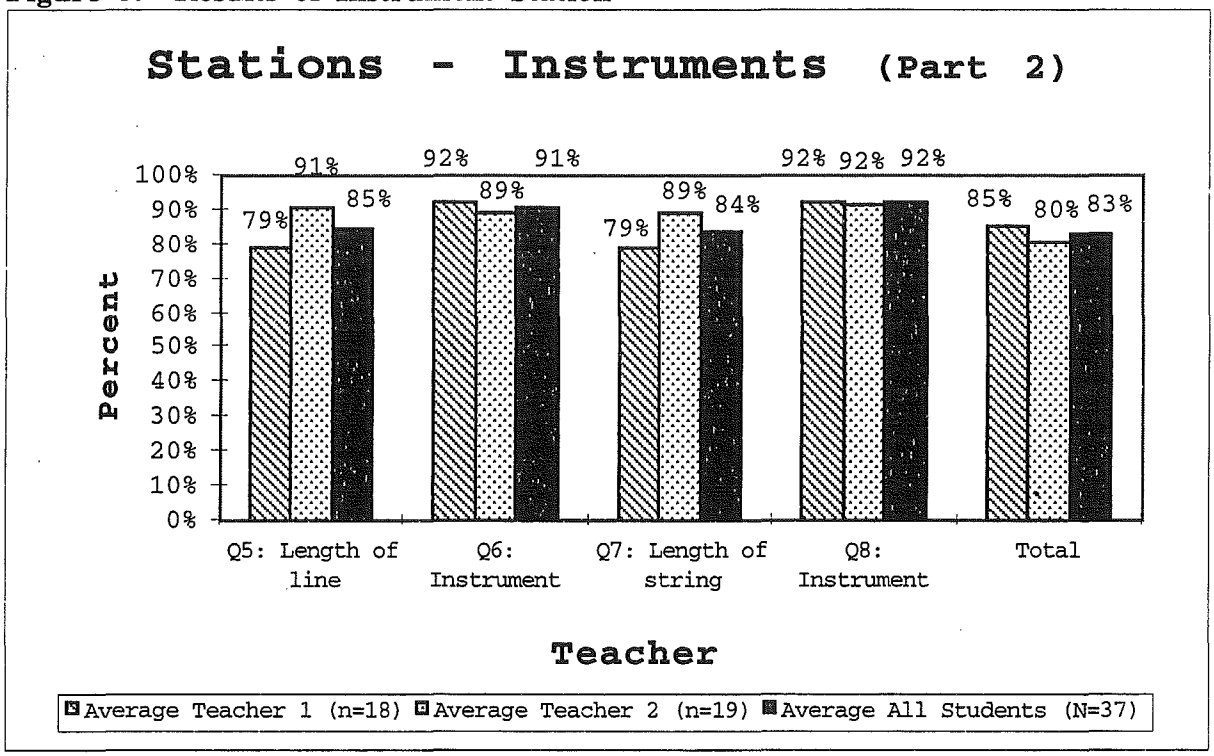
Figure 5: Results of Instrument Station



Question five an six concern the length of a line on a ball. Twenty-seven students (73%) answered accurately within plus or minus 1/2 inch, scoring three on a three-point scale or 100%. Four more measured accurately within plus or minus one inch scoring two points or 67%. Five gave other responses and one did not respond. Most (86%) used the right instrument scoring two on a two-point scale, although many did not know the correct name of the measuring tape.

Questions seven and eight pertain to the length of a piece of string. Twenty-seven students (73%) answered accurately within plus or minus 1/2 inch; three more came within and inch. Six gave other responses, scoring one or 33% on a three-point scale, and one did not respond. Most (86%) used the right instrument, a ruler or a measuring tape.

Figure 6: Results of Instrument Station



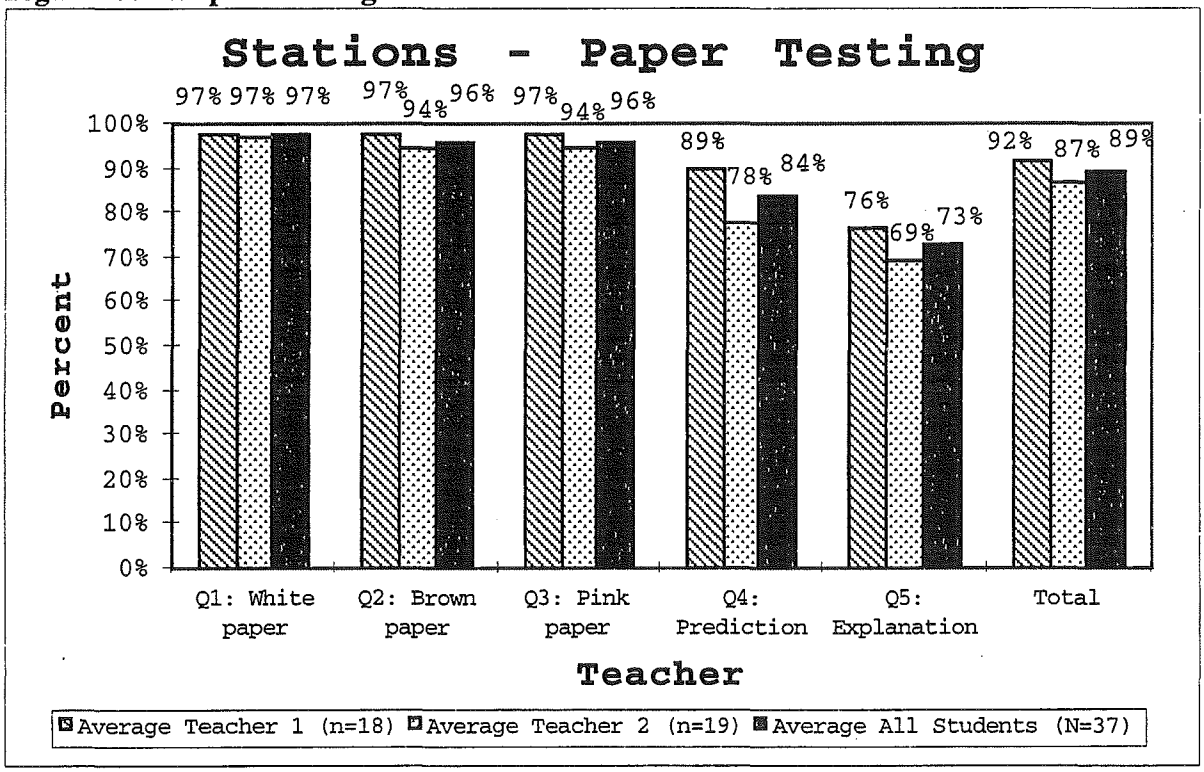
Station #4: Paper Testing

Figure 6 shows the results of the paper testing station. Overall, students earned 89% of the points possible at this station. For questions one, two, and three, students were asked to put a drop of water on a piece of white, brown and pink paper and observe the results. Each paper sample was different from the next in weight and texture. All but a couple of students offered at least one correct observation, scoring two points on a two-point scale.

Question four asked students to predict what would happen "if you could put a drop of water on the yellow paper in the plastic bag." The paper was similar in texture and weight to the white paper. Twenty-seven (73%) correctly predicted the result of dropping water onto the yellow paper. Two people left the item blank. The most common reason the remaining eight students made incorrect predictions was that they likened the yellow paper to the pink or brown paper. The group average was 84%.

Question five asked students to give the reasons they had for making their prediction. Twenty-two (59%) offered appropriate explanations for their correct predictions, scoring two points on a two-point scale. Ten students (27%) gave incorrect explanations; five left the item blank. The group average was 73%.

Figure 7: Paper Testing Results



Full Investigation: Sugar Testing

This test required students to design and conduct a complete experiment. Through this process students used a wide variety of skills found on the Exhibit Process Rubric such as making observations, conducting experiments, explaining results, making inferences and communicating findings. This test was completed by a sample of six students in each class for a total of twelve students. The students were asked to conduct a two-part experiment to answer the questions: "Do sugar cubes dissolve faster than loose sugar?" and "Does stirring make any difference in how fast the sugar cubes and loose sugar dissolve?"

The overall score for questions one and two are composite scores that include the scores for each of the questions parts; question one has four parts, question two has five. The total score is a composite of questions one and two. Comparisons were made between groups of students according to their teacher and according to their sex. Overall, no statistically significant differences were found.

Figure 7 shows scores for questions one and two and the total scores on the full investigation. Overall, students performed slightly better on the first part of the test than the second. The greatest factor contributing to this difference was the students' discussion of the results of the experiments. They were better able to discuss the results of the first experiment, scoring 81% as a group, than they were the results of the second experiment, scoring 60%. This may be due in part to the greater complexity of the second question. The first question asked "Which dissolves faster?" requiring the student to determine only whether loose sugar or a sugar cube dissolves faster. The second question asked "Does stirring make a difference?" This required the student to determine whether stirring has an effect and to state specifically what effect it has (with maximum points awarded to an answer indicating that both types dissolve faster than without stirring, and loose sugar still dissolves faster than a sugar cube). The total score for the students as a group was 59%.

Figure 8: Sugar Testing Results

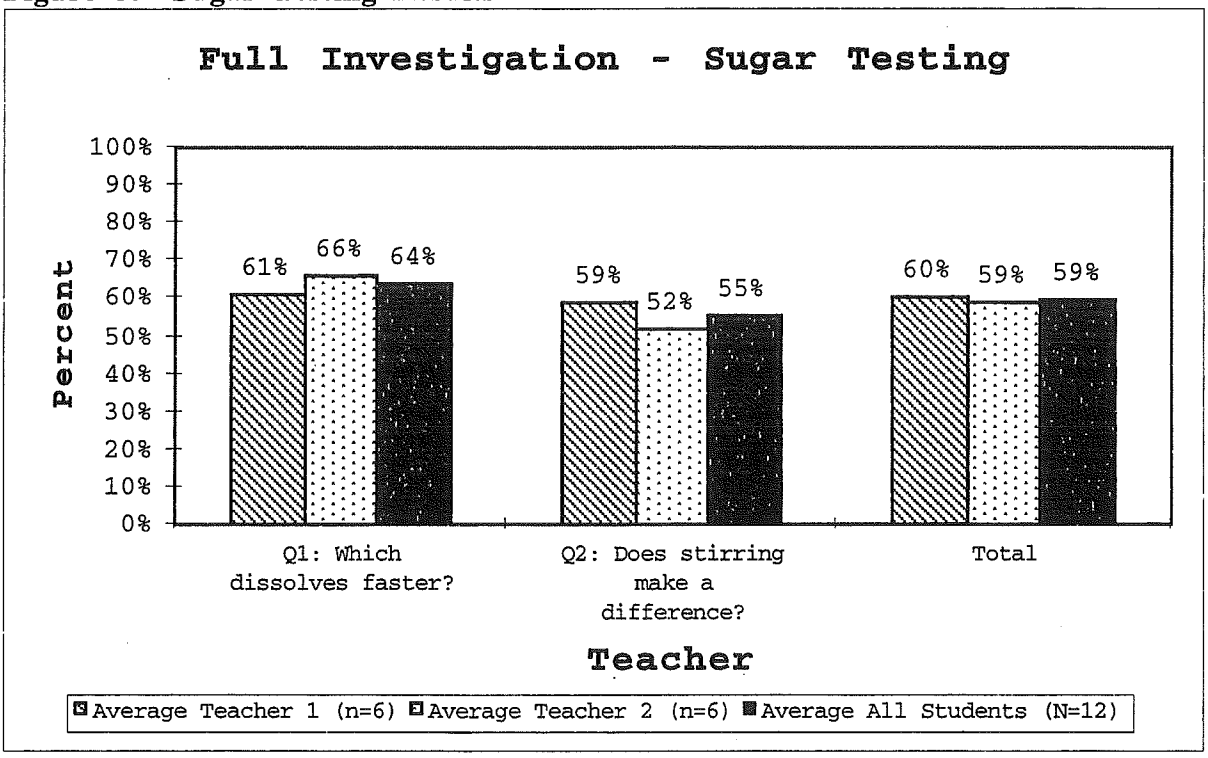


Figure 8 shows the component parts of question one, "Which dissolves faster?" In part 1.1, students were scored on whether or not they controlled the amount of water in each of their two cups. In order to score the maximum number of points, the students had to use equal amounts of water. Although several drew pictures indicating that approximately equal amounts of water were used, no student explicitly stated that he or she used equal amounts of water. Eleven of the twelve students made no mention of the amount of water (scoring 2 points on a 3-point scale, 67%); one used unequal amounts of water. The average score for the group was 64%.

Part 1.2 examined the method they used to measure the amount of time it took for each type of sugar to dissolve. The students could earn the maximum number of points by using a clock or a stopwatch. Of the twelve students, nine suggested only a visual comparison or made no mention of whether the two types of sugar were put into the water at the same time (scoring 4 on a 6-point scale, 67%). Only one made an explicit visual side-by-side comparison (5 of 6 points); none timed the experiment. The average score was 60%.

Part 1.3 scores students according to the number of times the experiment was conducted. All twelve students conducted the investigation only once (scoring 1 on a 2-point scale, 50%). The maximum number of points (2) would have been awarded if the experiment had been conducted more than once.

Part 1.4 asks students to discuss the results of the investigation. As a group, the average score was 81%. Eight of twelve, or two-thirds of the students correctly answered that loose sugar dissolves faster than sugar cubes (scoring 3 on a 3-point scale, 100%). One said that the sugar cube dissolved faster; three offered wrong or unrelated answers.

Figure 9: Sugar Testing Results - Question 1

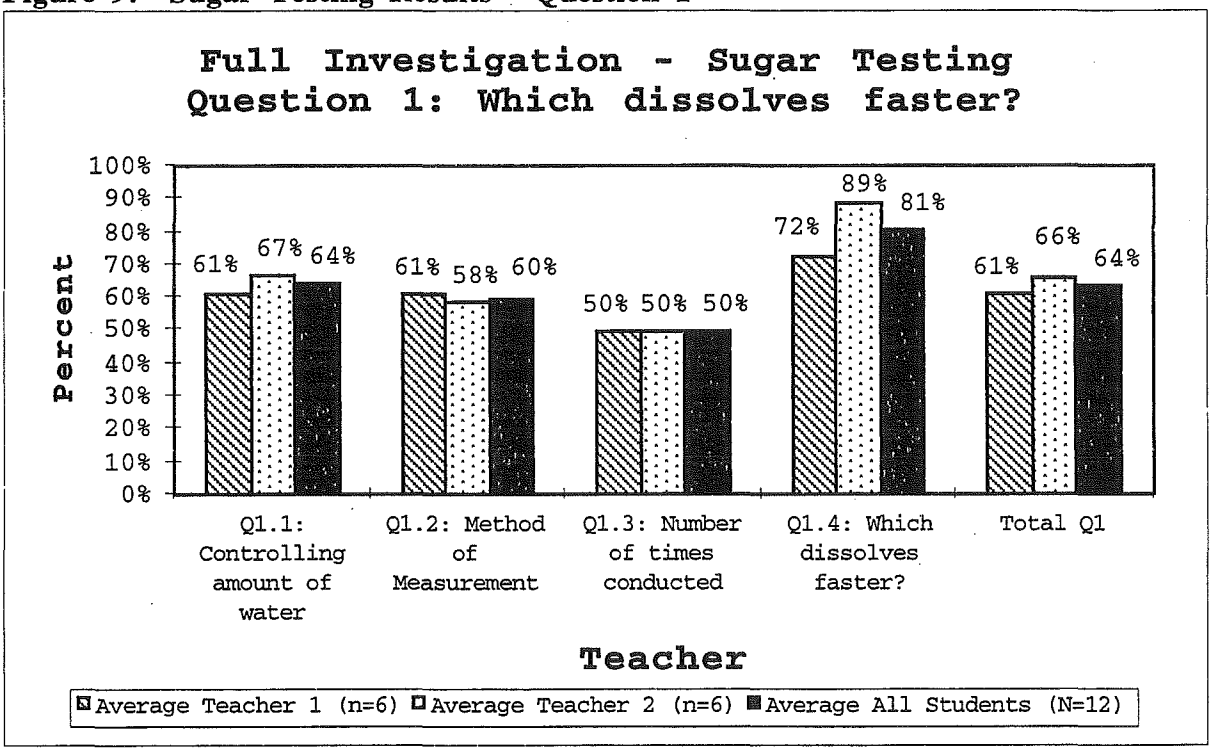


Figure 9 shows the component parts of question two, "Does stirring make a difference?" Part 2.1 scores students on whether they controlled the amount of water they used in their two cups. Again, eleven of the twelve students made no mention of the amount of water (scoring 2 points on a 3-point scale, 67%); one did not respond. The average score for the group was 61%.

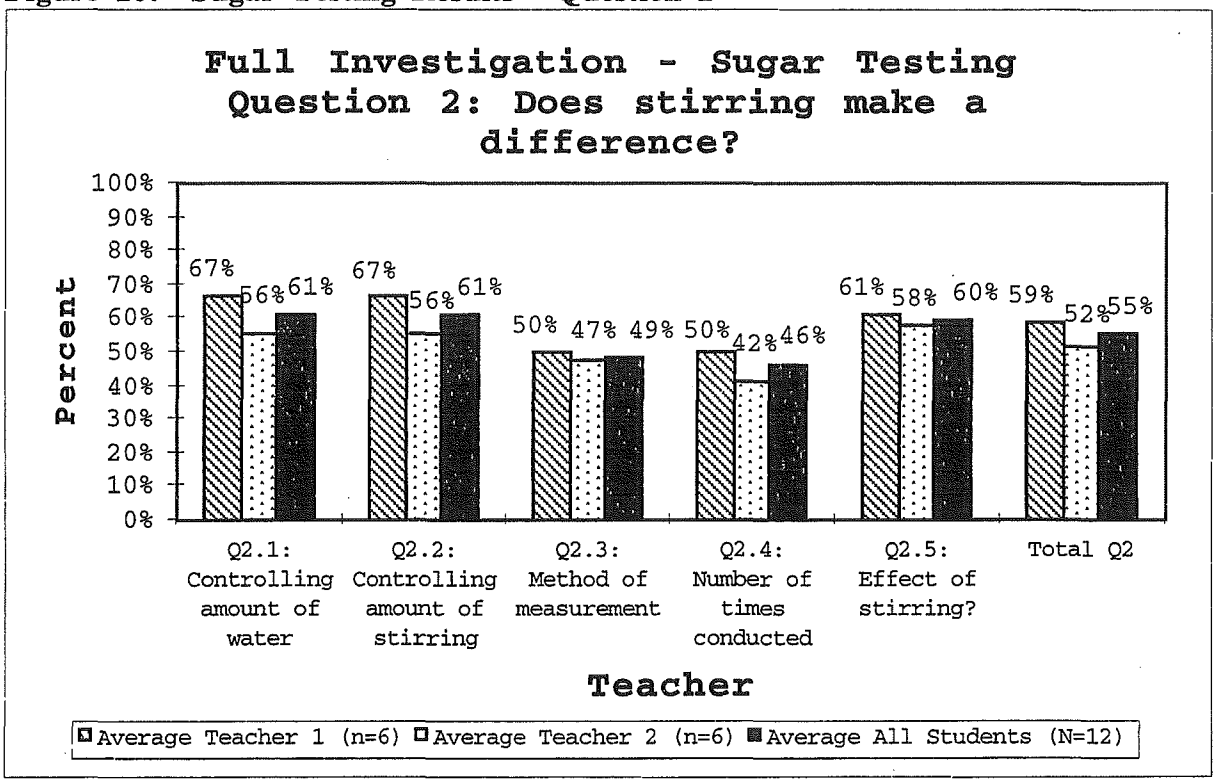
In part 2.2, students were scored on whether they controlled the amount of stirring they did in each of the two cups. To score the maximum number of points (3), the students would have explicitly stated that they stirred each equally. Eleven of the 12 students made no mention of the amount of stirring (scoring 2 points, 67%); one did not respond. The average was 61%.

Part 2.3 scores students on the method they used to measure the amount of time it took for each type of sugar to dissolve when it was stirred. This question saw a greater diversity of response types. The maximum number of points (6) was awarded if the student used a clock or stopwatch; one student did time the experiment and earned six points on a 6-point scale (100%). Five of 12 suggested a visual comparison or made no mention of whether the two types of sugar were put into the water at the same time (scoring 4 points of 6 possible, 67%). One student said that one type of sugar dissolves faster when stirred than when it was not stirred but made no comparison between the types of sugar (3 points, 50%); two came to the correct conclusion but did not mention stirring (2 points, 33%); two gave incorrect responses (1 point, 17%); one did not respond. The group average was 49%.

Part 2.4 scores students according to the number of times the experiment was conducted. The maximum number of points (2) would have been awarded if the experiment had been conducted more than once. Eleven of the 12 students conducted the investigation only once (scoring 1 point, 50%); one student did not respond. The group average was 46%.

Part 2.5 asks students to discuss the results of the investigation. This question also saw a diversity of response types. Two students scored the maximum number of points (6) by responding that both types of sugar dissolve faster when stirred and that the loose sugar still dissolves faster than the sugar cube. Two noted that loose sugar dissolves faster than sugar cubes when stirred (5 points). Three students saw that stirring makes a difference and made a general statement about what they saw (for example, that both dissolve faster). These students scored four points. Two students said that one type of sugar dissolves faster when stirred compared to when it is not stirred, but made no comparison between the types of sugar (3 points). One student came to the correct conclusion but did not mention stirring (2 points); one gave an incorrect response (1 point); one did not answer. The average score for the group was 60%.

Figure 10: Sugar Testing Results - Question 2



VII. Concluding Remarks

As a result of this project, staff at the MMS have established a process for defining and assessing what students learn as they develop exhibits. In the coming year, staff will continue to refine this process as they expand it to all grade levels and use the rubric as a tool for planning instruction and communicating with parents. In addition, the results of the hands-on testing conducted as part of this project provide information for staff to use as they continue to refine the exhibit development learning model. The implications of this project for future use of the exhibit development learning model are described below.

Implications for Strengthening the Exhibit Development Learning Model

The process of developing an assessment system for the exhibit development model has illuminated aspects of the model that need clarification. Future use of the model could be strengthened by clarifying issues such as the following:

- How do staff distinquish what gets taught and assessed in the exhibit development process from what gets taught in other areas of the curriculum? A strength of the model is that it provides a framework for interdisciplinary instruction, but this also makes it difficult to differentiate between what students learn by developing exhibits and what they learn through other classroom activities. There may not be one right answer to this question, but in order for the model to be used consistently across classrooms, the staff need a shared understanding of how the model will be used at this school.

- What are the distinguishing characteristics of the model, the practices that must occur in order for the model to be implemented effectively? The project results indicate that classroom teachers vary in how they integrate the exhibit process with the academic content they are required to cover. Some variation is necessary given the different needs of students and the different academic requirements at each grade level. In order for the model to have substance, however, there must be a core of activities and skills that distinguish the model from other instructional practices. Some possible distinctions may be that exhibit development always involves the use of real objects and always requires students to explain what they have done and learned to other people.

The hands-on tests administered in the project identified several specific exhibit process skills that students still need to develop. To strengthen their use of the model, staff may want to adjust their instruction to address these issues:

- Observe. According to the exhibit development rubric, proficiency is defined as making accurate, orderly observations. The results of this performance test, however, indicate that many students need development on this skill. Students were able to make simple straightforward observations, but rich, detailed descriptions of observations were rare. For example, on the popcorn station students watched a phenomenon with many different events to observe. When asked to describe what they observed, the majority of students only wrote down one or two observations with very minimal descriptions despite the fact there were at least six different events to observe. These results imply that students should be encouraged to make more careful, detailed observations.

•Design a Test / Conduct a Test. Designing and conducting a test is a complex set of skills and the results of the "Sugar Test" indicate that students need work in this area. One specific skill that needs work is learning how to conduct a "fair" test. On the sugar test, a fair test would involve using equal amounts of water to determine if loose sugar or cube sugar dissolves faster. One should also use equal amounts of sugar and put both types of sugar in the water for equal amount of time. None of the students who took the test explicitly controlled all of these variables. Students should be taught to carefully control variables to insure all tests they conduct are fair.

Not only is it important to use equal amounts of water, sugar and time, but it is also important to systematically measure each of these variables. Students were supplied with a stopwatch to measure the time and a measuring cup to measure the amount of water, but students very rarely used these instruments. They should be encouraged to use more precise measurement techniques.

Another important aspect of conducting a test is repeating the test more than once to make sure the results are consistent. None of students conducted the sugar test more than once.

•Explain. Students were also able to provide simple, straightforward explanations, but detailed explanations were lacking in students responses. Students were most likely to merely provide one explanation, rather than elaborate on their explanations. Students should be encouraged to write more detailed, thorough explanations.

Implications for Future Evaluation of the Model

Next steps for evaluating the exhibit development learning model are as follows:

•Develop rubrics for other grade levels. This task was already started at the summer workshops in August.

•Decide how often the hands-on testing should be used and which grade levels should be involved. The relationship between these tests and student ratings on the rubric also needs to be determined. For example, will results of the hands-on test influence teacher ratings of certain skills on the rubric or will the hands-on test results be reported separately?

•To determine the effectiveness of the exhibit development learning model in comparison to traditional instructional practices, administer the hands-on tests to a comparable group of students at a school that uses traditional instructional practices. In order for the comparison to be meaningful, the comparison school must have similar instructional goals.

•This project focused on the exhibit development learning model but there are other characteristics of the MMS that could be the focus of future evaluation efforts. For example, future studies could assess the effectiveness of other MMS components such as the emphasis on parent involvement, students working with adult mentors (either museum staff, community members, or parents), and students working cooperatively.

Appendix A:
Skills Students Learn During Exhibit Development

Skills Students Learn During Exhibit Development

As part of this project, two focus groups were held: one with teachers, and one with parents of students who had attended the school for several years. One of the focus group tasks was for parents and teachers to identify skills they believed students learned by developing exhibits. The results from each group are shown below.

Teacher Focus Group

Reading
Writing
Research
Speaking
Graphing
Listening
Observing
Quality Work
Use Scales
Gather / Use Materials
Technology
Cooperating
Sketching
Negotiating
Demonstrating
Performance Etiquette
Building / Construction
Graphic Arts
Planning / Organizing

Parent Focus Group

Learn How to Learn
Awareness of Different Cultures
Presentation Skills
Problem Solving
Critical Thinking
Teamwork
Confidence
Mazes
Saving, Observing, Collecting
Students become researchers
Ability to think about other possibilities
Reinforces asking a question and finding answers
Making connections, not learning things in isolation
Students gain the ability to wonder, raise questions, investigate, and classify

Appendix B:
Steps for Developing a Rubric to Assess Student Learning in the Exhibit Development Process

Steps for Developing a Rubric to Assess Student Learning in the Exhibit Development Process

This description is based on the process used to develop the rubric for the fourth grade at the Museum Magnet Elementary School.

1. Identify the skills that you think students learn through the instructional activity of exhibit development. You may want to identify separately the content knowledge learned. In the case of the Museum Magnet Elementary School, the project team wanted their assessment to be usable for any exhibit topic, and therefore decided not to include content knowledge. This does not mean that content is not important to them, it only means that content knowledge will not be assessed in this rubric. At the Museum Magnet Elementary School, the student report card has other areas that can be used to reflect students' content knowledge in areas such as mathematics or reading.

To generate ideas of skills covered in the exhibit process, consult publications such as the National Science Education Content Standards, the standards and performance packages that are part of the Minnesota Graduation Rule (or the standards of your state), and the learner outcomes developed by your district.

2. For each skill, you will describe three performance levels: needs development, proficient, and excellent. Begin by describing the performance of a student who is proficient in the skill in question. For example, the proficient level for the skill of "record results" would be "makes short notes with enough information to know what the notes refer to when read later." Proficiency is defined as the level of competence that is expected for most students at that grade level.

3. After describing what a proficient performance would look like, describe the performance of a student who needs development in this skill, and a student who demonstrates excellence in this skill. Performance at the excellent level should be beyond the expectation for most students at that grade level; generally, only about 5% of the students would be expected to demonstrate this level of the skill. Resources that are helpful in writing these descriptions include the assessment packages that are part of the Minnesota Graduation rule (or assessments related to your state standards), the National Science Education Content Standards, and materials provided as part of your district's work sampling system or standards assessments.

4. Try out the rubric, make notes about areas that need clarification, and refine the rubric. Teachers at the Museum Magnet Elementary School developed the following criteria to describe a good rubric: 1) The rubric should be written in clear language, without jargon; 2) the language should be understood by parents and students; and 3) the rubric should be useful as a tool for both planning and assessment.

On paper, the process of developing a rubric looks deceptively simple. In practice, you may find that you need to go back and forth between defining skills and describing what proficiency looks like for these skills. You may discover that some of the skills on your list overlap, or you may discover that some of the skills you listed are too complex and need to be broken up into separate skills. Using the rubric in the classroom will also help you identify areas that need further work; you should not expect that you will "get it right" the first time. Trying out the rubric will help you identify problems that you cannot identify just by looking at it.

Appendix C:
Hands-On Tests

Museum Magnet Performance Test

Name _____

Sex _____

Teacher _____

Date _____

STATION #1 - WIGWAG

Directions: Use the materials in front of you to answer the following questions.

1. **Lift each container one at a time. What do you notice about the containers?**

2. **Use a rubber band to put one of the containers on the end of the wigwag. Move the wigwag and notice how it moves. Now try it with the other containers.**

   What is the relationship between the weight of the containers and how the wigwag moves?

STATION #2 - POPCORN

Directions:
a. **Fill the plastic cup 3/4 full of water.**
b. **Drop 8-10 kernels of popcorn into the cup.**
c. **Drop 1 alka seltzer tablet into the cup.**
d. **Wait 30 seconds and observe.**

1. **Describe what you observed.**

_____

_____

_____

_____

2. **What explanation can you give for the motion of the popcorn?**

_____

_____

_____

_____

3. **Name <u>one</u> thing you would change to learn more about what is happening to the popcorn.**

_____

_____

_____

_____

******WHEN FINISHED, DUMP THE CUP INTO THE BUCKET******

STATION #3 - MEASURING TOOLS

Directions: Use the materials in front of you to answer the following questions.

1. What is the temperature of the room? _____ ^{0}F

   What tool did you use? _____

2. How long is the black line on the ball? _____ in

   What tool did you use? _____

3. How much water is in the bottle? _____ fl. oz.

   What tool did you use? _____

4. How long is the piece of string? _____ in

   What tool did you use? _____

STATION #4 - PAPER TESTING

Directions: a. Put <u>one</u> drop of water on each piece of paper.
b. Look at each drop with the magnifying glass.

1. **What happened to the drop of water on the:**

White Paper? _____

_____

Brown Paper? _____

_____

Pink Paper? _____

_____

2. **What do you think would happen if you could put a drop of water on the Yellow Paper in the plastic bag? (DO NOT OPEN BAG!)**

_____

_____

3. **Why do you think so?**

_____

_____

_____

_____

Conducting a Complete Experiment

Museum Magnet School

Name _____

Sex _____

Teacher _____

Date _____

Does stirring make any difference in how fast the sugar cubes and loose sugar dissolve?

Directions: Use the materials in front of you to answer the above question.

1. **Draw a picture of what you did.**

2. **Explain in words what you did.**

3. **Explain in words what you found out.**

Do sugar cubes dissolve faster than loose sugar?

Directions: Use the materials in front of you to answer the above question.

1. **Draw a picture of what you did.**

2. **Explain in words what you did.**

3. **Explain in words what you found out.**

Appendix D:
Tables of Hands-On Testing Results

Full Investigation - Sugar Testing
Results by Teacher
(N=12)

	Teacher 1 Average	Teacher 2 Average	Total Average	P-Value
Question 1: Which dissolves faster?	61%	66%	64%	0.65
1.1: Controlling amount of water	61%	67%	64%	0.34
1.2: Method of measurement	61%	58%	60%	0.83
1.3: Number of times investigation was conducted	50%	50%	50%	0.34
1.4: Result of investigation	72%	89%	81%	0.36
Question 2: Does stirring make a difference?	59%	52%	55%	0.55
2.1: Controlling amount of water	67%	56%	61%	.034
2.2: Controlling amount of stirring	67%	56%	61%	0.34
2.3: Method of measurement	50%	47%	49%	0.88
2.4: Number of times investigation was conducted	50%	42%	46%	0.34
2.5: Result of investigation	61%	58%	60%	0.89
Total	60%	59%	59%	0.85

Station #1 - WigWag
Results by Teacher
(N=37)

	Average Teacher 1	Average Teacher 2	Average All Students	P-Value
Q1: What do you notice about the containers?	79%	74%	77%	.53
Q2: What is the relationship between the weight of the containers and how the wigwag moves?	68%	59%	64%	.42
Total	74%	67%	70%	.37

Station #2 - Popcorn
Results by Teacher
(N=37)

	Average Teacher 1	Average Teacher 2	Average All Students	P-Value
Q1: Describe what you observed.	70%	61%	66%	.17
Q2: What explanation can you give for the motion of the popcorn?	42%	41%	41%	.82
Q3: Name one thing you would change to learn more about what is happening to the popcorn.	63%	58%	61%	.62
Total	58%	53%	56%	.39

Station #3 - Instruments
Results by Teacher
(N=37)

	Average Teacher 1	Average Teacher 2	Average All Students	P-Value
Q1: What is the temperature of the room?	77%	72%	75%	.64
Q2: What tool did you use?	95%	92%	93%	.66
Q3: How much water is in the bottle?	84%	52%	68%	.004*
Q4: What tool did you use?	84%	67%	76%	.05*
Q5: How long is the black line on the ball?	79%	91%	85%	.20
Q6: What tool did you use?	92%	89%	91%	.71
Q7: How long is the piece of string?	79%	89%	84%	.30
Q8: What tool did you use?	92%	92%	92%	.95
Total	85%	80%	83%	.43

Station #4 - Paper Testing
Results by Teacher
(N=37)

	Average Teacher 1	Average Teacher 2	Average All Students	P-Value
Q1: What happened to the drop of water on the white paper?	97%	97%	97%	.97
Q2: ...brown paper	97%	94%	96%	.53
Q3: ...pink paper	97%	94%	96%	.63
Q4: What do you think would happen if you could put a drop of water on the yellow paper in the plastic bag?	89%	78%	84%	.22
Q5: Why do you think so?	76%	69%	73%	.57
Total	92%	87%	89%	.40

All Stations
Results by Teacher
(N=37)

	Average Teacher 1	Average Teacher 2	Average All Students	P-Value
WigWag	74%	67%	70%	.37
Popcorn	58%	53%	56%	.39
Instruments	85%	80%	83%	.43
Paper Testing	92%	87%	89%	.40
Total	77%	72%	75%	.24

Full Investigation - Sugar Testing
Results by Sex of Students
(N=12)

	Female Average	Male Average	Total Average	P-Value
Question 1: Which dissolves faster?	65%	62%	64%	.75
1.1: Controlling amount of water	61%	67%	64%	.34
1.2: Method of measurement	58%	61%	60%	.83
1.3: Number of times investigation was conducted	50%	50%	50%	.34
1.4: Result of investigation	89%	72%	81%	.36
Question 2: Does stirring make a difference?	59%	52%	55%	.55
2.1: Controlling amount of water	67%	56%	61%	.34
2.2: Controlling amount of stirring	67%	56%	61%	.34
2.3: Method of measurement	53%	44%	49%	.64
2.4: Number of times investigation was conducted	50%	42%	46%	.34
2.5: Result of investigation	58%	61%	60%	.89
Total	62%	57%	59%	.60

Station #1 - WigWag
Results by Sex of Students
(N=37)

	Female Average	Male Average	Average All Students	P-Value
Q1: What do you notice about the containers?	79%	74%	77%	.53
Q2: What is the relationship between the weight of the containers and how the wigwag moves?	67%	61%	64%	.62
Total	73%	68%	70%	.51

Station #2 - Popcorn
Results by Sex of Students
(N=37)

	Female Average	Male Average	Average All Students	P-Value
Q1: Describe what you observed.	67%	64%	66%	.61
Q2: What explanation can you give for the motion of the popcorn?	42%	41%	41%	.82
Q3: Name one thing you would change to learn more about what is happening to the popcorn.	71%	50%	61%	.03*
Total	60%	52%	56%	.14

Station #3 - Instruments
Results by Sex of Students
(N=37)

	Female Average	Male Average	Average All Students	P-Value
Q1: What is the temperature of the room?	70%	80%	75%	.37
Q2: What tool did you use?	95%	92%	93%	.66
Q3: How much water is in the bottle?	72%	65%	68%	.55
Q4: What tool did you use?	76%	75%	76%	.89
Q5: How long is the black line on the ball?	81%	89%	85%	.38
Q6: What tool did you use?	97%	83%	91%	.10
Q7: How long is the piece of string?	84%	83%	84%	.93
Q8: What tool did you use?	97%	86%	92%	.12
Total	84%	82%	83%	.69

Station #4 - Paper Testing
Results by Sex of Student
(N=37)

	Female Average	Male Average	Average All Students	P-Value
Q1: What happened to the drop of water on the white paper?	97%	97%	97%	.97
Q2: ...brown paper	97%	94%	96%	.53
Q3: ...pink paper	95%	97%	96%	.68
Q4: What do you think would happen if you could put a drop of water on the yellow paper in the plastic bag?	84%	83%	84%	.93
Q5: Why do you think so?	79%	67%	73%	.31
Total	91%	88%	89%	.64

All Stations
Results by Sex of Student
(N=37)

	Female Average	Male Average	Average All Students	P-Value
WigWag	73%	68%	70%	.51
Popcorn	60%	52%	56%	.14
Instruments	84%	82%	83%	.69
Paper Testing	91%	88%	89%	.64
Total	77%	72%	75%	..32

Appendix E:
Scoring Guide for Hands-On Tests

Scoring Guide: Conducting a Complete Experiment

Question #1. Which dissolves faster?

1. Methods of Investigating

1.1 Controlling Amount of Water:

3 = equal amounts of water
- Student must be explicit. S/he must use words to describe the intentional action of controlling the amount of water in each of the cups.

2 = no mention of amount of water
- Student does not use words to indicate that s/he thought about controlling the amount of water, or uses only a picture to show that water amounts were approximately equal.

1 = unequal amounts of water
- Student explicitly states with words that unequal amounts of water were used.

0 = blank

1.2 Primary Method of Measurement:

6 = measured amount of time using clock or stopwatch
- The student must explain that s/he timed the experiment with a clock or stopwatch.

5 = visual side-by-side comparison
- The student must explicitly state that s/he put the loose sugar and the sugar cubes into the water at same time. Example: "First I put water into the cups. Then I put a sugar cube in the water at the same time I put the loose sugar into the water."

4 = visual comparison / no mention of measurement
- The student does not indicate that s/he measured the amount of time it took for the sugar to dissolve, and/or makes no mention of whether the two types of sugar were put into the water at the same time or not. Example: "I put the sugar into a cup of water and I also put a sugar cube into water."

3 = individual visual comparisons
- The student must explicitly state that s/he put one type of sugar into the water and then the other.

2 = used only one cup
- The student must explicitly state that s/he used one cup instead of two for the experiment. Example: "I put the sugar cube and sugar into the same cup with a little water and it melted."

1 = unclear
- The meaning of the student's response is unclear. Example: "I mixed a cube and sugar."

0 = blank

1.3 Number of times investigation was conducted:

2 = more than once
- Student explicitly states that s/he conducted the investigation more than once.

1 = only once
- Either the student explicitly states that s/he conducted the investigation only once, or this may be inferred from his/her description of the investigation.

0 = blank

1.4 Result of the investigation: Which type of sugar dissolves faster?

3 = loose sugar

2 = cube

1 = wrong answer or unrelated answer
- The meaning of the student's response is inaccurate, unrelated, or unclear. Example: "When you put sugar in and put in other glasses it will be clear."

0 = blank

Question #2. Does stirring make a difference?

2. Methods of Investigating

2.1 Controlling Amount of Water:

3 = equal amounts of water (must be explicit)

2 = no mention of amount of water

1 = unequal amounts of water (must be explicit)

0 = blank

2.2 Controlling the Amount of Stirring:

3 = equal amounts of stirring
- Student must be explicit. S/he must use words to describe the intentional action of controlling the amount of stirring used in each of the cups.

2 = no mention of equal stirring
- Student does not use words to indicate that s/he thought about controlling the amount of stirring applied to each cup. Example: "I poured the sugar pack in and stirred it then I put in a sugar cube and stirred it."

1 = unequal amounts of stirring
- Student explicitly states that s/he stirred the cups unequally.

0 = blank

2.3 Primary Method of Measurement:

6 = measured amount of time using clock or stopwatch

5 = visual side-by-side comparison (loose and cubes at same time; must be explicit)

4 = visual comparison / no mention of at the same time or not

3 = individual visual comparisons (one type of sugar then the other; must be explicit)

2 = used only one cup

1 = unclear

0 = blank

2.4 Number of Times Investigation was Conducted:

2 = more than once

1 = only once

0 = blank

2.5 Results of Investigation: Effect of Stirring?

6 = both types dissolve faster than not stirring, but loose still fastest
- Student may articulate his/her answer in different ways, but it must be clear that s/he came to the above conclusion. Examples: "I found out that it dissolves quicker when you stir it. It took 2 minutes 43 seconds for a sugar cube to dissolve while being stirred and it took 19 seconds for loose sugar to dissolve." "Stirring the loose sugar make it dissolve faster and when I stirred it with the cube in it only dissolved a little bit faster."

5 = loose sugar dissolves faster than cube when stirred
- Student correctly indicates which type of sugar dissolves faster, but does not indicate that stirring makes a difference for both types of sugar. Example: "The sugar cube didn't dissolve as fast as the loose sugar when I was stirring."

4 = stirring makes a difference; general statement only about why stirring matters
- Student must draw a correct conclusion about what difference stirring makes, and must explicitly state that stirring affects both types of sugar. Example: "I found out that they both go faster if you stir."

3 = mentions one type of sugar dissolves faster when stirred
- Student discusses the effect of stirring on one type of sugar and makes no comparison between the two types of sugar. Example: "I found out that the sugar cubes dissolve faster if you stir them."

2 = correct conclusion, doesn't mention stirring
- Student comes to the correct conclusion, but does not explicitly state that stirring makes a difference. Example: "I found out that the loose sugar melts faster than the sugar cube and the sugar cube melts lesser than the loose sugar."

1 = incorrect response
- The meaning of the student's response is inaccurate, unrelated, or unclear. Example: "The water looks like sticky water. It also looks like the water is with gas."

0 = blank

Scoring Guide: Museum Magnet Performance Test
Station #1 - WigWag

I. What do you notice?

3 = A=lightest, B= medium, C= heaviest

- Student must describe the weight of the containers, explicitly stating that A was the lightest, B was heavier, and C was the heaviest. Example: "A is very light, B is heavy, but C is the heaviest."

2 = containers have different weights

- Student makes an observation about the weight of the containers, but does not describe each of the containers and compare them to one another. Examples: "They get heavier." "C is more heavier." "Some of the containers weigh more than the other ones."

1 = incorrect observation

- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: "The first one has not a thing in it. I think B has a nickel in it. I think C has a penny in it." "One got nothing in it. The other one have something in it and B is the heaviest." "I think they are gonna fall."

0 = blank

II. Relationship between weight and wigwag?

3 = the heavier the container, the slower (or lower) it moves or vice versa

- Student must explicitly state that the heavier the container, the slower the wigwag moves. Also acceptable are observations that the heavier the container, the lower the wigwag moves. Examples: "The heavier the container, the lower the wigwag moves." "When I put A on it went fast. When I put B on it went slower. Then I put C on it hardly moved." "A is light and when I put it on the wigwag it barely moved. B made it bend and move a lot since it is heavy. Once again since C is the heaviest, it made the wigwag bend and move the most.

2 = wigwag moves differently with different containers (does not mention relationship)

- Student observes that the wigwag moves differently with containers of different weights, but does not comment on the systematic nature of the relationship of the movement to the differing weights of the containers. Examples: "The second container weighed more and was bending the wigwag." "When you put the containers on the wigwag A was lighter because it didn't have nothing in it and B, C was heavy so it made the wigwag on down." "The wigwag is kept moving with the two with the weight and A stopped sooner."

1 = incorrect response

- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: "If the container is light it goes really high. If the container is heavy it don't wiggle. And letter C bends really far." "C is the heaviest and A is the lightest. The same about each container is that A, B and C are letters of the alphabet." "If you put A on the wigwag it will move. If you put C on it won't be moving."

0 = blank

Station #2 - Popcorn

I. Observations

4 = three or more relevant observations
- Relevant observations include, for example: the alka seltzer dissolves, bubbles form, kernels go up, popcorn then sinks, etc.

3 = two relevant observations

2 = one relevant observation

1 = incorrect observation
- The meaning of the student's response is inaccurate, unrelated, or unclear. If there is one relevant observation and one incorrect this should be scored as a "2"; a score of "1" is for a student who only wrote a wrong observation.

0 = blank

II. Explanation

3 = explicit, complete explanation of phenomenon
- Student gives an explicit, complete explanation, i.e., compares weights, density, etc.

2 = reasonable explanation
- Student gives a reasonable but incomplete explanation, such as "the bubbles make the popcorn go up." Examples: "I think the bubbles from the alka seltzer made the popcorn move." "Cause most of the air was pushing up and then some of the popcorn went up and stuck on the top and some stuck on the bottom." "Air bubbles attached to the popcorn and rose."

1 = incorrect response
- Responses that only offer additional observations (not explanations), or responses that are inaccurate, unrelated, or unclear. Examples: "First it spun around and then it started rising. Only two pieces stayed on the bottom of the cup." "Some popcorn got held down with bubbles." "It start popping up to the top of the cup." "I think that the pressure probably pushed the popcorn up."

0 = blank

III. Changes

2 = modify one of the original conditions
- Student suggests a change or modification of one of the original conditions: liquid, popcorn, quantity, add something, use "popped" popcorn, etc. Examples: "What would happen if you put salt in?" "More of the alka seltzer." "I would put more kernels in the cup and maybe another alka seltzer."

1 = incorrect, irrelevant response
- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: "Find out why bubbles are forming on it." "Do not put the popcorn in." "Some float up to the top."

0 = blank

Station #3 - Instruments

Temp of Room:
3 = +/- 2° F
2 = +/- 4° F
1 = other (If other, did student use Celsius temp? 1=yes, 2=no)
0 = blank

Instrument
2 = thermometer
1 = other
0 = blank

Volume of Water:
3 = +/- 1 fl oz
2 = +/- 2 fl oz
1 = other
0 = blank

Instrument
2 = measuring cup
1 = other
0 = blank

Length of Line:
3 = +/- 1/2 inch
2 = +/- 1 inch
1 = other
0 = blank

Instrument
2 = ruler or tape measure
1 = other
0 = blank

Length of String:
3 = +/- 1/2 inch
2 = +/- 1 inch
1 = other
0 = blank

Instrument
2 = ruler or tape measure
1 = other
0 = blank

Note: In scoring the "instrument used" questions, the scorer may accept as correct answers inexact spellings, inexact names (such as "measuring line" and "measuring bucket"), and pictures.

Station #4 - Paper Testing

White paper

2 = correct observation
- Student offers one or several correct observations such as nothing happened, the drop stayed on surface, did not soak in, etc. Examples: "The drop just stayed the same. It didn't soak into the paper." "It formed a droplet." "It just stayed there."

1 = incorrect
- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: "The water on the paper did get ripe." "It growed a little and moved."

0 = blank

Brown paper

2 = correct
- Student offers one or several correct observations such as the water soaked in, the paper absorbs the water, etc. Examples: "The drop spreaded and soaked through the paper." "It spread in a circle." "It spread a lot and darkened the paper." "It soaked right through."

1 = incorrect
- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: Examples: "The brown paper got water on it." "The drop on the brown paper went on the paper."

0 = blank

Pink paper

2 = correct
- Student offers one or several correct observations such as the water soaked in, the paper absorbs the water, etc. "The water soaked through and became colored pink. The paper also lost its texture." "The water took the color out of the paper." "It got bigger than it was at first."

1 = incorrect
- The meaning of the student's response is inaccurate, unrelated, or unclear.

0 = blank

Prediction

2 = correct (e.g.
- Student offers one or several correct predictions such as nothing would happen, the drop would stay on paper, not soak in, like white paper. Examples: "I think it would stay there." "It will do the same as the white paper."

1 = incorrect
- The meaning of the student's response is inaccurate, unrelated, or unclear. Examples: "It would soak through." "It will get water onto it." "It would roll off."

0 = blank

Explanation

2 = correct
- Student offers one or several correct explanations such as the paper looks like white paper, thick paper won't absorb, etc. Should be interpreted in the context of the previous question on answer given for prediction. Examples: ("It do the same as the white paper) because it feels the same." "(It would roll off.) Yellow paper is like white paper." "Because the surface of the yellow paper and the white paper seem to be the same."

1 = incorrect
- The meaning of the student's response is inaccurate, unrelated, or unclear. Should be interpreted in the context of the answer given for the previous question on prediction. Examples: "(It will soak through) because it is a lighter color." "(It will turn colors) because the paper is yellow and when I put a drop on the pink it turned colors." "I think it will make it coal."

0 = blank